



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# Deletion of the Polycomb-Group Protein EZH2 Leads to Compromised Self-Renewal and Differentiation Defects in Human Embryonic Stem Cells

### Citation for published version:

Collinson, A, Collier, AJ, Morgan, NP, Sienerth, AR, Chandra, T, Andrews, S & Rugg-Gunn, PJ 2016, 'Deletion of the Polycomb-Group Protein EZH2 Leads to Compromised Self-Renewal and Differentiation Defects in Human Embryonic Stem Cells', *Cell Reports*, vol. 17, no. 10, pp. 2700-2714.  
<https://doi.org/10.1016/j.celrep.2016.11.032>

### Digital Object Identifier (DOI):

[10.1016/j.celrep.2016.11.032](https://doi.org/10.1016/j.celrep.2016.11.032)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Peer reviewed version

### Published In:

Cell Reports

### Publisher Rights Statement:

This article is available under the terms of the Creative Commons Attribution License (CC BY). You may copy and distribute the article, create extracts, abstracts and new works from the article, alter and revise the article, text or data mine the article and otherwise reuse the article commercially (including reuse and/or resale of the article) without permission from Elsevier. You must give appropriate credit to the original work, together with a link to the formal publication through the relevant DOI and a link to the Creative Commons user license above. You must indicate if any changes are made but not in any way that suggests the licensor endorses you or your use of the work.

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# **Deletion of the Polycomb-group protein EZH2 leads to compromised self-renewal and differentiation defects in human embryonic stem cells**

Adam Collinson<sup>1</sup>, Amanda J. Collier<sup>1,2</sup>, Natasha P. Morgan<sup>1</sup>, Arnold R. Sienerth<sup>1</sup>, Tamir Chandra<sup>1</sup>, Simon Andrews<sup>3</sup> and Peter J. Rugg-Gunn<sup>1,2,4,\*</sup>

<sup>1</sup>Epigenetics Programme, The Babraham Institute, Cambridge, CB22 3AT, UK.

<sup>2</sup>Wellcome Trust – Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge, CB2 1QR, UK.

<sup>3</sup>Bioinformatics Group, The Babraham Institute, Cambridge, CB22 3AT, UK.

<sup>4</sup>Centre for Trophoblast Research, University of Cambridge, Cambridge, CB2 3EG, UK.

\*Lead Author (peter.rugg-gunn@babraham.ac.uk).

**Abstract**

Through the histone methyltransferase EZH2, the Polycomb complex PRC2 mediates H3K27me3 and is associated with transcriptional repression. PRC2 regulates cell-fate decisions in model organisms, however, its role in regulating cell differentiation during human embryogenesis is unknown. Here, we report the characterization of *EZH2*-deficient human embryonic stem cells (hESC). H3K27me3 was lost upon *EZH2* deletion, identifying an essential requirement for EZH2 in methylating H3K27 in hESC, in contrast to its non-essential role in mouse ESC. Developmental regulators were derepressed in *EZH2*-deficient hESC, and single-cell analysis revealed an unexpected acquisition of lineage-restricted transcriptional programmes. *EZH2*-deficient hESC show strongly reduced self-renewal and proliferation, thereby identifying a more severe phenotype compared to mouse ESC. *EZH2*-deficient hESC can initiate differentiation towards developmental lineages, however, they cannot fully differentiate into mature specialised tissues. Thus, *EZH2* is required for stable ESC self-renewal, regulation of transcriptional programmes and for late-stage differentiation in this model of early human development.

**Keywords:** pluripotency, differentiation, epigenetics, histone methylation

## Introduction

Polycomb-group (PcG) proteins are epigenetic repressors of transcriptional programmes and maintain cellular identity during development, differentiation and disease (Di Croce and Helin, 2013; Pasini and Di Croce, 2016; Pietersen and van Lohuizen, 2008; Schuettengruber and Cavalli, 2009; Surface et al., 2010). PcG proteins form two well-characterized and biochemically distinct chromatin-modifying complexes that are termed Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2). PRC1 catalyzes histone H2A lysine 119 ubiquitination through the activity of the E3 ligases RING1A and RING1B (Muller and Verrijzer, 2009; Wang et al., 2004). PRC2 is composed of the core proteins EZH2, EED and SUZ12, together with RBAP46/48 and several other accessory subunits, and is responsible for catalyzing di- and trimethylation on histone H3 lysine 27 (H3K27me<sub>2/3</sub>) (Cao et al., 2002; Czermin et al., 2002; Kuzmichev et al., 2002; Margueron and Reinberg, 2011; Muller et al., 2002). EZH2 is a SET-domain containing histone methyltransferase and is the catalytic subunit of PRC2. EED and SUZ12 are required for substrate recognition, complex stability and for promoting the enzymatic activity of EZH2 (Cao and Zhang, 2004; Nekrasov et al., 2005; Pasini et al., 2004; Tie et al., 2007).

Genome-wide studies in mouse and human embryonic stem cells (ESC) have shown that PRC2 and H3K27me<sub>3</sub> occupy the promoters of many developmental regulators that are important for cell differentiation and lineage specification (Azuara et al., 2006; Bernstein et al., 2006; Boyer et al., 2006; Bracken et al., 2006; Lee et al., 2006; Mikkelsen et al., 2007; Pan et al., 2007; Zhao et al., 2007). This distribution of chromatin marks led to the concept that PRC2 may contribute to the maintenance of pluripotency by keeping developmental regulators transcriptionally repressed, whilst enabling the genes to be rapidly activated upon suitable differentiation cues and stimuli. Despite a central position within the regulatory framework, however, PRC2 is dispensable for the maintenance of undifferentiated mouse ESC, as the deletion of PRC2 components, including *Ezh2*, has little effect on their morphology, self-

renewal or proliferation, although a subset of PRC2 target genes are modestly derepressed (Chamberlain et al., 2008; Leeb et al., 2010; Pasini et al., 2007; Riising et al., 2014; Shen et al., 2008). H3K27me3 levels are globally reduced in *Ezh2*-deficient mouse ESC, however, developmental regulators retain H3K27me3 at their gene promoters and are transcriptionally repressed (Shen et al., 2008). In this context, the *Ezh2* homologue, *Ezh1*, forms a noncanonical PRC2 complex that is able to trimethylate H3K27 at target gene promoters and maintains transcriptional repression through methylation-dependent and potentially methylation-independent pathways (Margueron et al., 2008; Shen et al., 2008).

PRC2-deficiency has a more significant impact on mouse ESC upon their differentiation, with defects in the repression of pluripotency networks and in the failure to fully activate differentiation transcriptional programmes. This aberrant gene regulation results in impaired differentiation and proliferation (Chamberlain et al., 2008; Pasini et al., 2007; Shen et al., 2008). Further underscoring the critical role of PRC2 in directing differentiation programmes, all three PRC2 core components (*Ezh2*, *Eed* and *Suz12*) are essential for early mouse development, as loss of function mutant embryos initiate but fail to complete gastrulation and die between embryonic days 7 and 9 (Faust et al., 1995; O'Carroll et al., 2001; Pasini et al., 2004). The mutant phenotype is associated with mis-expression of lineage-specifying genes, decreased cell proliferation and an increased level of apoptosis (Pasini et al., 2004).

The well-conserved binding profiles of PRC2 components and H3K27me3 in human ESC (hESC) at the promoters of developmental regulators raises the possibility that PRC2 may also have an important role in controlling hESC pluripotency and differentiation (Gifford et al., 2013; Ku et al., 2008; Pan et al., 2007; Zhao et al., 2007). Moreover, coordinated changes to the epigenome, including H3K27me3 localization, occur upon differentiation of hESC and are thought to be essential for lineage specification and memory of cellular identity, as they are in *Drosophila* and the mouse (Gifford et al., 2013; Xie et al., 2013). However, no functional

studies of PRC2 in hESC regulation and early stage differentiation have been reported to date. Human pluripotent cells represent a unique model in which to study human development, and provide a platform for producing a source of differentiated cells relevant for basic and applied research. In addition, mouse and human ESC are known to represent different pluripotent states and may therefore rely on different epigenetic pathways to confer their ability to self-renew and differentiate (Nichols and Smith, 2009; Rossant, 2015). Understanding the key epigenetic mechanisms that underpin hESC is therefore a priority.

Here, we report the generation and characterization of *EZH2*-deficient hESC. Our findings demonstrate that EZH2 is required to maintain the transcriptional repression of developmental regulators and for cells to undergo late-stage cell differentiation, thereby revealing the broad conservation of PRC2 function in this model of early human development. We also identify unexpected human-specific differences such as the essential requirement in hESC for EZH2 to maintain PRC2 stability and retain promoter-localised H3K27me3, which is in contrast to its non-essential role in mouse ESC. In addition, self-renewal and proliferation are also perturbed to a greater extent in *EZH2*-deficient hESC, as compared to *Ezh2*-deficient mouse ESC. Our study therefore provides a comprehensive characterization of PRC2 function in hESC, thereby providing a new platform to investigate the role of histone methylation in regulating the genome during human development and stem cell differentiation.

## Results

### Targeted deletion of *EZH2* in hESC

To investigate the role of *EZH2* in human pluripotency and differentiation, we used CRISPR/Cas9 to disrupt *EZH2* in hESC. A guide RNA (gRNA) designed to target an early exon within all known *EZH2* isoforms was nucleofected with *Cas9* into the H9 hESC line (Figures 1A and S1A). Individual colonies were isolated, expanded and analysed by Sanger DNA sequencing. The efficiency of disrupting the target sequence within the *EZH2* coding region

was high, with ~35% clonal lines containing a mutation on one allele (*EZH2*<sup>-/+</sup>). However, no homozygous cell lines were obtained out of 110 screened lines. This result provided a first indication that *EZH2*-deficient hESC may be compromised relative to *EZH2*-containing cells when plated as single cells at clonal density. To overcome this apparent defect, we introduced a doxycycline (DOX)-inducible *EZH2* transgene using piggyBac transposition into an *EZH2*<sup>-/+</sup> line and re-targeted the cells with *EZH2* gRNA and *Cas9* in the presence of DOX. Using this strategy we obtained several *EZH2* homozygous lines (*EZH2*<sup>-/-</sup>; Figures 1A and S1B,C). Once the *EZH2*<sup>-/-</sup> lines were isolated and established, they could be maintained without DOX-induced *EZH2* expression. Although we did not detect any indication that the DOX-inducible plasmid was leaky in the absence of DOX, to rule out the possibility of low level *EZH2* expression, we transiently transfected *EZH2*<sup>-/-</sup> ESC with piggyBac transposase and obtained stable *EZH2*<sup>-/-</sup> lines with all copies of the *EZH2* transgene removed (Figures S2A,B).

RNA expression analysis confirmed that *EZH2* transcripts were lower in *EZH2*<sup>-/-</sup> ESC compared to parental *EZH2*<sup>+/+</sup> and *EZH2*<sup>-/+</sup> lines (Figures 1B and S2C). Moreover, *EZH2* protein was undetectable by Western blot and by immunofluorescent microscopy using two different antibodies raised against N- and C-terminal epitopes of *EZH2* (Figures 1C,D and S2D-F). The disruption of *EZH2* was accompanied by the loss of other PRC2 proteins, *SUZ12* and *EED*, despite the presence of unchanged levels of *SUZ12* and *EED* transcripts in *EZH2*<sup>-/-</sup> ESC (Figures 1B,D). This finding unexpectedly contrasts with *Ezh2*-deficient mouse ESC where *Suz12* and *Eed* levels are unchanged due to the ability of *Ezh1* to form noncanonical PRC2 (Shen et al., 2008), but is consistent with *Suz12*-deficient and *Eed*-deficient mouse ESC in which PRC2 components are unstable outside of the complex (Pasini et al., 2007; Pasini et al., 2004). In hESC, therefore, *EZH1* is unable to form noncanonical PRC2 despite being present

(Figure S2G). *EZH1* transcript and protein levels were largely unchanged upon *EZH2* deletion (Figure S2G). Immunofluorescent microscopy revealed that the loss of *EZH2* led to the reduction of H3K27me3 and H3K27me2 to background levels, and to the partial reduction of H3K27me1 (Figure 1E). Applying DOX to induce ectopic *EZH2* expression in *EZH2*<sup>-/-</sup> ESC restored EZH2 and led to the stabilization of SUZ12 and EED proteins, and to the re-establishment of global H3K27 methylation (cells designated herein as *EZH2*<sup>-/-</sup> +*EZH2*; Figures 1B-E).

### Loss of promoter-localised H3K27me3 in *EZH2*-deficient hESC

To characterise the molecular phenotype of *EZH2*-deficient hESC, we profiled genome-wide histone methylation by native chromatin immunoprecipitation combined with high-throughput sequencing (ChIP-Seq). Quantitative trend plots of normalized ChIP-Seq reads revealed a complete loss of H3K27me3 at all gene promoters in *EZH2*<sup>-/-</sup> ESC (Figure 2A). This finding contrasts with the retention of H3K27me3 at the promoters of developmental regulators in *Ezh2*-deficient mouse ESC (Shen et al., 2008). Confirming the result in hESC, scatter plot analysis of ~2000 promoters that have high levels of H3K27me3 in *EZH2*<sup>+/+</sup> ESC (H3K27me3<sup>WT</sup>) revealed a loss of H3K27me3 in *EZH2*<sup>-/-</sup> ESC (Figure 2B). The majority of H3K27me3<sup>WT</sup> promoters have histone H3 lysine 4 trimethylation (H3K4me3) in hESC (Pan et al., 2007; Zhao et al., 2007), and H3K4me3 levels were largely unaffected by *EZH2* disruption (Figure 2B). ChIP-Seq tracks for two example loci, *HOXB* and *HOXD*, illustrate the loss of H3K27me3 across the domains in *EZH2*<sup>-/-</sup> ESC (Figure 2C). Comparison of H3K27me3 between *EZH2*<sup>+/+</sup> and *EZH2*<sup>-/-</sup> +*EZH2* cells revealed highly similar profiles, demonstrating that histone patterns are appropriately re-established upon EZH2 restoration (Figures 2A-C). Interestingly, there was a modest increase in histone H3 lysine 27 acetylation (H3K27ac) levels at H3K27me3<sup>WT</sup> promoters in *EZH2*<sup>-/-</sup> ESC, supporting a potential antagonism between



H3K27 acetylation and trimethylation that has also been observed in other contexts (Ferrari et al., 2014; Gehani et al., 2010; Jung et al., 2010; Pasini et al., 2010; Schoenfelder et al., 2015) (Figure S3A). Lastly, an alternative ChIP-Seq analysis strategy of genome binning confirmed a loss of H3K27me3 sequencing reads across the genome of *EZH2*<sup>-/-</sup> ESC, further reinforcing the key finding that H3K27me3 levels are depleted upon deletion of *EZH2* in hESC (Figure S3B). Together, these results demonstrate that EZH2 is the main functional H3K27me2/3 methyltransferase in hESC.

### ***EZH2*-deficiency causes transcriptional derepression of key developmental genes**

We next performed RNA sequencing (RNA-Seq) to investigate the impact of loss of EZH2 and associated H3K27me3 on gene expression. The assays were carried out on samples that were flow-sorted using the hESC cell-surface marker SSEA4 to ensure that we compared between equivalent cell populations (Figure S4A). The majority of genes were not altered transcriptionally by *EZH2* disruption, but 911 genes were significantly upregulated and 282 genes were significantly downregulated in *EZH2*<sup>-/-</sup> ESC compared to *EZH2*<sup>-/-</sup> +*EZH2* ESC ( $p < 0.05$ ; Figures 3A and S4B). Gene ontology (GO) analysis of the upregulated gene set identified categories associated with developmental and cellular differentiation, including pattern specification, embryonic morphogenesis and tissue formation (Figure 3B). The upregulated group was significantly enriched for genes with EZH2 and H3K27me3 occupancy in *EZH2*<sup>+/+</sup> ESC and are thus expected to be sensitive to *EZH2* disruption (Figures 3C,D). Notably, not all genes with EZH2-bound promoters were mis-regulated, suggesting that secondary events may be required for this group of genes to initiate transcriptional change in response to the loss of EZH2. In addition, a subset of upregulated genes (enriched for signalling and adhesion GO terms) are likely to be regulated indirectly as they are not PRC2 targets in hESC. No GO categories were significantly enriched in the downregulated gene set,

although of note, the top category was associated with the regulation of M-phase and indicate the decreased transcription of genes typically expressed in mitosis (Figure 3B).

We examined transcriptional changes in a set of ~100 classical developmental regulators that have strong EZH2 promoter occupancy, including genes within the FOX, GATA, LHX, T-box and SOX families (Lee et al., 2006). A clear pattern emerges from this analysis: nearly all genes within this class of developmental regulator showed transcriptional derepression in the absence of *EZH2* (Figures 3E and S4C). Gene derepression did not become more prevalent upon continued passaging, and therefore the *EZH2*<sup>-/-</sup> ESC retained a similar transcriptional profile over time. ChIP-Seq and RNA-Seq tracks for several example genes, including *SOX17*, *GATA4*, *T* and *TBX3*, illustrate the absence of promoter H3K27me3 and an associated increased transcript level in *EZH2*<sup>-/-</sup> ESC (Figure 3F). We further showed that shRNA-mediated depletion of *EZH2* causes derepression of developmental regulators in two additional human pluripotent stem cell lines (WiBR3 and FiPS; Figure S4D). We conclude that the deletion of *EZH2* in hESC leads to the loss of H3K27me3 and to the transcriptional derepression of genes that encode developmental regulators, thereby positioning EZH2 as a key factor in controlling the transcriptome of human cell types during early development.

### Single-cell transcriptional analysis reveals gene mis-regulation profiles

To investigate more precisely the transcriptional mis-regulation and cell-to-cell variability in response to *EZH2* deficiency, we performed single-cell RNA-Seq on individual SSEA4-positive, flow-sorted *EZH2*<sup>-/-</sup> and *EZH2*<sup>+/+</sup> ESC. The results show that a subset of *EZH2*<sup>-/-</sup> ESC strongly upregulated EZH2-target genes, but individual genes are not robustly derepressed in most cells examined (Figure 4A). Unexpectedly, clustering of the data suggested that gene derepression occurs predominately within discrete transcriptional programmes, such that individual *EZH2*<sup>-/-</sup> ESC have upregulated multiple genes associated with a particular cell

lineage but rarely show strong signatures derived from several cell lineages (Figure 4B). In particular, *EZH2*<sup>-/-</sup> ESC show lineage biases towards endoderm and mesoderm, but not to ectoderm. This response may be constrained by the hESC culture environment due to the activities of FGF and Activin/Nodal signalling within the media, which are known to promote endoderm and mesoderm specification and suppress ectoderm differentiation (Pauklin and Vallier, 2015). Together, these results reveal that the depletion of EZH2 does not cause global EZH2-target gene derepression in all ESC, as might be predicted from cell population studies. Rather, its loss leads to the mis-regulation of subsets of genes and to the acquisition of lineage-restricted transcriptional programmes.

### ***EZH2*-disruption causes self-renewal and proliferation defects in hESC**

Undifferentiated *EZH2*-deficient ESC could be maintained in culture for >50 passages, however, their growth and morphology were severely compromised compared to control lines. *EZH2*<sup>-/-</sup> colonies were highly variable in appearance with an increased prevalence of flatter cells that are characteristic of spontaneous differentiation (Figure 5A). To compare directly the ability of each hESC line to self-renew, we plated an equal number of SSEA4-positive flow sorted cells at clonal density and after seven days we counted the number of colonies that were positive for the undifferentiated hESC marker alkaline phosphatase (AP). We observed a ~40% reduction in colony number in *EZH2*<sup>-/-</sup> ESC compared to control ESC (Figure 5B).

Diminished colony formation was due to both compromised self-renewal and impaired proliferation. Categorising *EZH2*<sup>-/-</sup> ESC colonies based on AP activity patterns revealed a ~50% reduction in the proportion of undifferentiated colonies in *EZH2*<sup>-/-</sup> ESC compared to control ESC, and an associated increase in the proportion of colonies with a mixed or fully

differentiated phenotype (Figure 5C). This was accompanied by a decrease in the proportion of *EZH2*<sup>-/-</sup> ESC colonies that are entirely OCT4-positive, with an associated increase in the proportion of colonies that are positive for SOX17, a marker of early differentiated cells (Figure 5D). The proportion of undifferentiated and differentiated cells within the *EZH2*<sup>-/-</sup> ESC cultures was unchanged over passage, and was re-established after plating of purified undifferentiated *EZH2*<sup>-/-</sup> ESC, further highlighting the unstable nature of these cells.

Cell counts over four passages revealed a ~50% reduction in cell number in *EZH2*<sup>-/-</sup> ESC, revealing that proliferation is significantly reduced in the absence of *EZH2* (Figure 5E). The mitotic index, as determined by the proportion of histone H3 serine 10 phosphorylation (H3S10ph) positive cells, was significantly lower in *EZH2*<sup>-/-</sup> ESC compared to control ESC (Figure 5F). The reduction in mitotic cells within *EZH2*<sup>-/-</sup> cultures is in agreement with our RNA-Seq results, which identify the transcriptional downregulation of genes associated with M-phase (Figure 3B). Of further relevance to this defect is that several negative regulators of the cell cycle, such as *CDKN2A* (encoding p16INK4A and p14ARF) and *CDKN2B* (encoding p15INK4B), were transcriptionally derepressed in *EZH2*-deficient ESC (Figure S5). These findings are consistent with previous studies in other cell types that have identified a role for EZH2 in controlling the transcription of cell cycle regulators (Bracken et al., 2007; Bracken et al., 2003; Pasini et al., 2004; Sauvageau and Sauvageau, 2010; Varambally et al., 2002). Together, the results demonstrate that *EZH2*-deficient hESC are strongly compromised in their ability to self-renew and proliferate, thereby identifying a more severe phenotype compared to mouse ESC that are deficient for *Ezh2* and other PRC2 proteins.

### **Differentiation defects in *EZH2*-deficient hESC**

We next investigated the impact of *EZH2* deletion on the ability of hESC to differentiate correctly. We injected each hESC line into the kidney capsule of three immunocompromised mice to test for teratoma formation. Control hESC lines produced teratomas consisting of mature cell types derived from all three germ lineages. By contrast, *EZH2*<sup>-/-</sup> ESC failed to produce teratomas in two mice and generated a very small mass in one mouse, which consisted of a restricted set of cell types, including immature adipocytes and epithelial cells (Figures 6A and S6A,B). Although the DOX-inducible *EZH2* transgene could partially rescue the *EZH2*<sup>-/-</sup> phenotype, we noticed that the teratomas formed from the *EZH2*<sup>-/-</sup> +*EZH2* ESC were smaller and displayed different morphology compared to the *EZH2*<sup>+/+</sup> and *EZH2*<sup>-/+</sup> ESC teratomas (Figure 6A and S6A,B). The difference is likely to be because the cells were not provided with DOX once they were injected *in situ*, and therefore EZH2 levels would be lost gradually over several days. Further investigation of *EZH2* function in late-stage cell differentiation revealed that very few *EZH2*<sup>-/-</sup> ESC survived after five days of retinoic acid-mediated differentiation *in vitro*, compared to *EZH2*<sup>+/+</sup> and *EZH2*<sup>-/+</sup> ESC (Figure 6B). Restoration of *EZH2* with the DOX-inducible transgene partially rescued the defect, although we noticed that the transgene was silenced at the later stages of cell differentiation, thereby hindering a full rescue. Together, these results lead us to conclude that hESC require *EZH2* to form late-stage differentiated cell types.

We next studied the early stages of ESC differentiation. PRC2 loss of function mutant embryos initiate but fail to complete gastrulation (Faust et al., 1995; O'Carroll et al., 2001; Pasini et al., 2004), however, a detailed examination of PRC2-deficient mouse or human ESC differentiation towards early developmental progenitors using defined conditions has not been reported. To investigate these developmental events, we initiated directed and separate differentiation towards endoderm, mesoderm and ectoderm progenitors, using defined

conditions and lineage-specific markers (Supplemental Experimental Procedures). We observed that cell number declined sharply during endoderm and mesoderm differentiation of *EZH2*<sup>-/-</sup> ESC, such that very few cells remained by the end of their differentiation protocols (Figures 6C,D). In contrast, cell number was maintained in *EZH2*<sup>+/+</sup> ESC during differentiation. We examined the cells at a mid-stage timepoint as this allowed us to obtain sufficient cells for analysis. Quantitative analysis of lineage-specific, cell-surface markers using flow cytometry showed that *EZH2*<sup>-/-</sup> ESC were capable of differentiating into early stage endoderm (defined as KIT<sup>+</sup>/CXCR4<sup>+</sup>) (Nostro et al., 2011) and mesoderm (defined as PDGFRα<sup>+</sup>/KDR<sup>+</sup>) (Kattman et al., 2011) and, surprisingly, formed cell populations at these time points that were more uniform in marker expression than achieved upon differentiation of *EZH2*<sup>+/+</sup> ESC (Figures 6C,D). To ascertain whether a pre-existing subset of endoderm progenitors were responsible for generating endoderm cells in *EZH2*<sup>-/-</sup> cultures, we used flow cytometry to separate KIT<sup>+</sup>/CXCR4<sup>+</sup> (endoderm primed) and KIT<sup>-</sup>/CXCR4<sup>-</sup> (not endoderm primed) *EZH2*<sup>-/-</sup> populations and subjected the cells to endoderm differentiation. Flow cytometry analysis showed that KIT<sup>-</sup>/CXCR4<sup>-</sup> were highly efficient in generating endoderm, thereby demonstrating the ability of *EZH2*-deficient ESC to respond to appropriate differentiation cues and initiate early stage differentiation (Figure S6C).

Upon cell differentiation, pluripotency factors *POU5F1*, *NANOG* and *SOX2* were downregulated to a similar extent in *EZH2*<sup>-/-</sup> ESC compared to *EZH2*<sup>+/+</sup> ESC (Figures 6E,F). This finding is in contrast to PRC2-deficient mouse ESC, which exhibit a defect in silencing pluripotency networks during cell differentiation (Pasini et al., 2007; Shen et al., 2008). Consistent with our RNA-Seq data, genes associated with endoderm and mesoderm differentiation were detected at higher levels at day 0 in *EZH2*<sup>-/-</sup> ESC compared to *EZH2*<sup>+/+</sup>

ESC, and the transcript level of these genes increased further during differentiation, confirming their ability to undergo early stage differentiation (Figures 6E,F). Ectopic expression of lineage-restricted genes occurred during *EZH2*<sup>-/-</sup> ESC differentiation, suggesting a failure to repress alternate transcriptional programmes (Figures 6E,F).

*EZH2*-deficient cells were also able to generate early ectoderm cells (defined as CD56<sup>+</sup>/CD326<sup>-</sup>) (Gifford et al., 2013), however, cell number during differentiation and the efficiency of differentiation were significantly reduced compared to *EZH2*<sup>+/+</sup> ESC (Figure 6G). Mis-regulation of lineage-restricted genes was observed upon ectoderm differentiation, thereby identifying a requirement for EZH2 in regulating appropriate gene expression in the early stages of ectoderm specification (Figure 6H).

Taken together, these results demonstrate that EZH2 is not required for the initial phase of hESC differentiation, but is required for the robust generation of mature cell types that are produced in the later stages of differentiation *in vitro* or in teratoma assays.

## Discussion

PcG-proteins are essential regulators of cell fate decisions and transcriptional programmes during the development of several species, including *Drosophila* and the mouse. Here, we show that this important function is also required in humans during the establishment of early developmental cell types that arise upon ESC differentiation (Figure 7). Furthermore, loss of EZH2 function in undifferentiated hESC led to the transcriptional derepression of ~900 genes including many important developmental regulators, thereby positioning EZH2 as a key factor in controlling the transcriptome of human cell types during early development. Interestingly, not all PRC2-target genes were mis-regulated, suggesting that redundant modes of transcriptional repression are in place, or that additional cues (such as transcription factor binding) are required to fully activate those genes. In addition to developmental factors, we

also detected an increased expression of cell-cycle regulators in the absence of *EZH2*, including *CDKN2A* (encoding p16INK4A and p14ARF) and *CDKN2B* (encoding p15INK4B). Given the close association between cell-cycle control and cell differentiation in hESC (Pauklin and Vallier, 2013; Gonzales et al., 2015; Pauklin et al., 2016), it is likely that both processes contribute to the phenotype of *EZH2*-deficient cells. For example, an upregulation of p16INK4A would inhibit CDK4/6, which, in turn, would lead to an increase in Activin/Nodal activity (Pauklin and Vallier, 2013). This signalling change would promote the transcriptional programmes of endoderm and mesoderm lineages, and suppress ectoderm differentiation. We speculate, therefore, that the increased expression of endoderm and mesoderm genes in *EZH2*-deficient hESC is caused jointly by the removal of repressive H3K27me3 marks, by altered signalling activities that are mediated by cell-cycle machinery, and by the cell culture environment. Notably, our transcriptional results are consistent with a recent study that reported derepression of a subset of PRC2-target genes in *Ezh2*-deficient mouse epiblast tissue, thereby underscoring the relevance of our observations to pluripotent cells *in vivo* (Zylicz et al., 2015). Importantly, our analysis of individual cells further revealed that misregulation of genes tended to occur in a coordinated manner within lineage-restricted transcriptional programmes, rather than a haphazard derepression of all PRC2-target genes as might be predicted from global cell population analysis. These results suggest the presence of feedback mechanisms that are able to promote or repress alternative cell fates during the early phases of differentiation. An exciting set of future studies will be to model and investigate the mechanisms responsible for this feedback. Our findings also raise broader questions about how cells are committed to a particular lineage during differentiation. Purifying live hESC populations that are in different transcriptional states and challenging the cells to functional assays should begin to unravel the complexities of cell fate commitment during human development.



Despite extensive conservation in their functions, differences exist between mouse and human ESC that lack *EZH2*; self-renewal, morphology and proliferation are seemingly perturbed to a greater extent in human *EZH2*<sup>-/-</sup> ESC compared to mouse *Ezh2*<sup>-/-</sup> ESC (Shen et al., 2008). One potential explanation is that differences in PRC2 protein stability or function could contribute to the distinct mouse and human ESC phenotypes. For example, Eed and Suz12 levels are unaffected by the loss of Ezh2 in mouse ESC, potentially due to a partial compensation by Ezh1 (Shen et al., 2008). In contrast, we show here that depletion of EZH2 in hESC results in loss of EED and SUZ12, despite the presence of EZH1. Interestingly, Ezh1 cannot compensate for the absence of Ezh2 during mouse ESC differentiation or embryo gastrulation (O'Carroll et al., 2001; Shen et al., 2008). We speculate there is a context-dependent role for Ezh1 and that the compensatory function diminishes as cells enter the post-implantation phase of development, which could partially explain the apparent inability of EZH1 to fulfil a compensatory role in *EZH2*-deficient hESC. In addition, Ezh1 is able to repress gene transcription through methylation-independent mechanisms in somatic cells, potentially via chromatin compaction (Margueron et al., 2008). It will therefore be interesting in future studies to more precisely define the functional interplay between EZH1 and EZH2 in early human developmental cell types.

A second potential explanation for the distinct phenotypes is that mouse and human ESC are known to represent different pluripotent states, with hESC considered to be primed for differentiation (Nichols and Smith, 2009; Rossant, 2015). The *EZH2*-deficient phenotype may therefore manifest differently depending on cell state; a concept recently proposed for *DNMT1*-depleted ESC (Liao et al., 2015). It will be important in future studies to test this hypothesis by investigating the role of PRC2 in human 'naïve' pluripotent cells, which are reported to be more similar to mouse ESC (Manor et al., 2015). Furthermore, it is interesting to consider that PRC2 may contribute to the balance required for primed-state pluripotency by enabling low-level expression of lineage-specifying developmental regulators whilst

constraining their levels so that they do not overwhelm the maintenance of the undifferentiated state. Given that the ectopic expression of several EZH2 target genes, such as *SOX17* and *GATA6*, can induce the differentiation of hESC (Seguin et al., 2008; Wamaitha et al., 2015), it is plausible that derepression of these and other developmental regulators in the absence of *EZH2* results in a shift towards an increased level of spontaneous differentiation that is observed in *EZH2*-deficient hESC. Thus, our analysis of PcG function in hESC should lead to a better understanding of the processes that regulate lineage priming and cell fate commitment, and inform similar events that occur in other species and cell types.

Genome-wide mapping has revealed that dynamic changes in epigenetic marks, including H3K27me3 localization, occur upon hESC differentiation (Gifford et al., 2013; Xie et al., 2013). A prevailing model proposes that this epigenetic reconfiguration is required to coordinate transcriptional programmes and provide a memory of cell identity. We have now tested this model and we show that EZH2 is not required for the initial phase of hESC differentiation as ectoderm, mesoderm and endoderm germ lineages can form in the absence of EZH2, however, the mutant cells mis-express lineage-specific genes, are unstable and are gradually lost over the differentiation timecourse. Interestingly, and in contrast to PRC2-deficient mouse ESC (Pasini et al., 2007; Shen et al., 2008), pluripotency genes were downregulated appropriately upon differentiation of *EZH2*-deficient hESC, suggesting that these genes are silenced by PRC2-independent pathways. The observed differentiation defects and reduction in cell survival are therefore unlikely to be caused by aberrant expression of pluripotency factors, but rather by mis-expression of lineage-specifying genes and cell-cycle regulators. Finally, although the *EZH2*-deficient hESC were unable to form mature cell types, the rescue of early differentiation defects by conditionally restoring *EZH2* levels should enable the role of PRC2 to be investigated during late-stage *in vitro* differentiation. As the *EZH2*<sup>-/-</sup> +*EZH2* cells could not fully recapitulate the parental wild-type cells in the teratoma and the RA-

differentiation experiments, alternative conditional systems might be better suited for the investigation of *EZH2* in late-stage cell differentiation. Interestingly, the results from our teratoma experiments suggest that once the *EZH2*<sup>-/-</sup> ESC have overcome the initial early-stage differentiation barrier (enabled by residual *EZH2*), they are able to specialize along certain tissue lineages. The differences in morphology and tissue composition of the teratomas are presumably, to some extent, a reflection of what differentiation pathways are accessible to *EZH2*-deficient cells. Future studies using alternative conditional strategies and more precise differentiation systems should lead to a better understanding of epigenetic modifiers in the generation of specialized cell types. Artificially controlling *EZH2* levels may also have useful practical applications in producing desired cell types, as has been demonstrated recently to boost production of beta cell progenitors (Xu et al., 2014).

Taken together, our study provides a comprehensive examination of *EZH2* function in hESC pluripotency and differentiation. Of note is that PRC2 mediates the self-renewal and differentiation of adult stem cells and cancer stem cells (Sauvageau and Sauvageau, 2010). Our findings therefore not only reveal the role of epigenetic modifiers and associated histone marks in regulating the genome during early human development, but also establish general principles that can be applicable to stem cells involved in homeostasis and disease.

## Experimental Procedures

### Cell culture

HESC (H9/WA09, obtained from WiCell; WIBR3, kindly provided by Rudolph Jaenisch; FiPS, kindly provided by Austin Smith) were cultured at 37°C in 5% CO<sub>2</sub> in air on CF1 irradiated mouse embryonic fibroblasts (MEF) in Advanced DMEM containing 20% Knockout Serum Replacement supplemented with 2mM L-Glutamine, 0.1mM  $\beta$ -mercaptoethanol, 1x Penicillin/Streptomycin, 1x Non-Essential Amino Acids (all from Thermo Fisher Scientific) and 4ng/ml FGF2 (WT-MRC Cambridge Stem Cell Institute). Where indicated, DOX was added at 1 $\mu$ g/ml. For feeder-free culture, ESC were transferred onto Vitronectin matrix in TeSR-E8 media (StemCell Technologies). Authentication of the hESC was achieved by confirmation of expression of pluripotency gene and protein markers. Cells were routinely verified as mycoplasma-free using a PCR-based assay. Additional cell culture materials and methods are detailed in Supplemental Experimental Procedures.

### Targeted deletion of *EZH2*

*EZH2* gRNA (CCGCTTCTGCTGTGCCCTTATC) was designed using <http://crispr.mit.edu> (Hsu et al., 2013). The gRNA sequence was incorporated into the U6 target gRNA expression vector (Mali et al., 2013) and synthesised as a gBlock by Integrated DNA Technologies. The *EZH2* gRNA gBlock was sub-cloned into pCR2.1-TOPO (Thermo Fisher Scientific) and verified by sequencing. HESC were dissociated into single cells using Accutase (Thermo Fisher Scientific). H9 ESC (2 million) were nucleofected with 5 $\mu$ g pCas9\_GFP (Addgene plasmid # 44719) and 5 $\mu$ g *EZH2* gRNA expression vector. After 48h, 10,000 GFP-positive single cells were isolated by FACS and seeded onto MEF in a 10cm tissue culture dish in ESC media supplemented with 10 $\mu$ M Rho Kinase inhibitor (Cell Guidance Systems) for the first 24h. Individual clones were picked and expanded in 24 well plates. Mutations were validated by DNA sequencing of TOPO

cloned PCR products. As a check for specificity, ten predicted off-target gRNA sites within genes were tested and verified to contain unmodified sequences.

### Plasmid constructs

To construct PB-TET-EZH2-ires-mCherry plasmid, the *EZH2* coding sequence was amplified using primers EZH2\_attb\_F and EZH2\_attb\_R and sub-cloned into PB-TET-ires-mCherry plasmid. To generate *EZH2*<sup>-/-</sup> +*EZH2* ESC, *EZH2*<sup>-/+</sup> ESC were lipofected with 1µg PB-TET-EZH2-ires-mCherry, 1µg pCAG-rtTA-Puro and 2µg pCyL43 (Wang et al., 2008) followed by selection with 1µg/ml Puromycin.

To remove PB-TET-EZH2-ires-mCherry, DOX-induced *EZH2*<sup>-/-</sup> +*EZH2* ESC were nucleofected with 5µg pCMV-hyPBase (Yusa et al., 2011) and 1µg Turbo-GFP (Lonza). After 48h, 10,000 GFP / mCherry double-positive single cells were isolated by FACS and seeded onto MEF in a 10cm tissue culture dish in ESC media supplemented with 10µM Rho Kinase inhibitor for the first 24h. Individual clones were picked and expanded in 24-well plates. DNA was genotyped using mCherry\_Geno and TET-Prom\_Geno primers to confirm removal of PB-TET-EZH2-ires-mCherry.

### Statistics

For Figure 3D, the data are significantly departed from normality ( $p < 0.05$ ; D'Agostino-Pearson omnibus normality test) and the variance is different between the groups ( $p < 0.05$ ; Brown-Forsythe test) therefore a non-parametric test was used. For statistical analysis of data within Figures 5 and 6, the scatter of the data lead us to assume that the samples comes from a normally distributed population and that the variability between the groups is about the same, therefore parametric tests were used.

### Data availability

Sequencing data have been deposited in Gene Expression Omnibus with accession number GSE76626.

### **Author contributions**

A.C. and P.J.R.-G. designed the study, interpreted the results and wrote the manuscript. A.C. generated all cell lines and performed all experiments. A.J.C. carried out the ectoderm differentiation experiments. N.P.M. generated several ChIP-Seq libraries. A.R.S. performed cell line characterization. T.C. assisted with generating the single-cell RNA-Seq libraries. S.A. analysed single-cell RNA-Seq data. P.J.R.-G. conceived and supervised the project, performed experiments and analysed data. We consider A.J.C. and N.P.M. to have equal contributions.

### **Acknowledgements**

We thank members of several Babraham Institute Facilities including Rachael Walker and Arthur Davies in Flow Cytometry, Kristina Tabbada in Sequencing, Anne Segonds-Pichon and Felix Krueger in Bioinformatics, and Simon Walker and Hanneke Okkenhaug in Imaging. We are indebted to William Mansfield at the WT-MRC Cambridge Stem Cell Institute, and Fernando Constantino-Casas and Madeline Fordham at the Department of Veterinary Medicine at the University of Cambridge, for performing the teratoma experiments and histology, respectively. We are grateful to Rudolph Jaenisch at the Whitehead Institute for Biomedical Research for providing WIBR3 cells, Austin Smith at the WT-MRC Cambridge Stem Cell Institute for providing FiPS cells, Manousos Koutsourakis and Bill Skarnes at the Wellcome Trust Sanger Institute for helpful advice with the use of CRISPR/Cas9 in hESC, Kiran Musunuru at Harvard University for pCas9\_GFP, and Anastasiya Sybirna for constructing the *EZH2* shRNA plasmid. We thank Sarah Elderkin, Wolf Reik and members of the Rugg-Gunn group for helpful discussions and comments on the manuscript. P.J.R.-G. is

supported by the Wellcome Trust (WT093736) and BBSRC (BBS/E/B/000C0402). A.C. and A.J.C. are supported by MRC DTG Studentships (MR/J003808/1).

## References

- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merckenschlager, M., *et al.* (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8, 532-538.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.
- Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., Ziller, M., Croft, G.F., Amoroso, M.W., Oakley, D.H., *et al.* (2011). Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* 144, 439-452.
- Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K., *et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349-353.
- Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* 20, 1123-1136.
- Bracken, A.P., Kleine-Kohlbrecher, D., Dietrich, N., Pasini, D., Gargiulo, G., Beekman, C., Theilgaard-Monch, K., Minucci, S., Porse, B.T., Marine, J.C., *et al.* (2007). The Polycomb group proteins bind throughout the INK4A-ARF locus and are disassociated in senescent cells. *Genes Dev* 21, 525-530.
- Bracken, A.P., Pasini, D., Capra, M., Prosperini, E., Colli, E., and Helin, K. (2003). EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer. *EMBO J* 22, 5323-5335.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298, 1039-1043.



- Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol Cell* 15, 57-67.
- Chamberlain, S.J., Yee, D., and Magnuson, T. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* 26, 1496-1505.
- Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* 111, 185-196.
- Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. *Nat Struct Mol Biol* 20, 1147-1155.
- Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. *Development* 121, 273-285.
- Ferrari, K.J., Scelfo, A., Jammula, S., Cuomo, A., Barozzi, I., Stutzer, A., Fischle, W., Bonaldi, T., and Pasini, D. (2014). Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Mol Cell* 53, 49-62.
- Gehani, S.S., Agrawal-Singh, S., Dietrich, N., Christophersen, N.S., Helin, K., and Hansen, K. (2010). Polycomb group protein displacement and gene activation through MSK-dependent H3K27me3S28 phosphorylation. *Mol Cell* 39, 886-900.
- Gifford, C.A., Ziller, M.J., Gu, H., Trapnell, C., Donaghey, J., Tsankov, A., Shalek, A.K., Kelley, D.R., Shishkin, A.A., Issner, R., *et al.* (2013). Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell* 153, 1149-1163.
- Gonzales, K.A., Liang, H., Lim, Y.S., Chan, Y.S., Yeo, J.C., Tan, C.P., Gao, B., Le, B., Tan, Z.Y., Low, K.Y., Liou, Y.C., Bard, F., and Ng, H.H. (2015). Deterministic restriction on pluripotent state dissolution by cell-cycle pathways. *Cell* 162:564-579.
- Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., *et al.* (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 31, 827-832.

- Jung, H.R., Pasini, D., Helin, K., and Jensen, O.N. (2010). Quantitative mass spectrometry of histones H3.2 and H3.3 in Suz12-deficient mouse embryonic stem cells reveals distinct, dynamic post-translational modifications at Lys-27 and Lys-36. *Mol Cell Proteomics* 9, 838-850.
- Kattman, S.J., Witty, A.D., Gagliardi, M., Dubois, N.C., Niapour, M., Hotta, A., Ellis, J., and Keller, G. (2011). Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell Stem Cell* 8, 228-240.
- Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S., *et al.* (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* 4, e1000242.
- Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev* 16, 2893-2905.
- Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125, 301-313.
- Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev* 24, 265-276.
- Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., *et al.* (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat Genet* 47, 469-478.
- Loh, K.M., and Lim, B. (2011). A precarious balance: pluripotency factors as lineage specifiers. *Cell Stem Cell* 8, 363-369.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823-826.

- Manor, Y.S., Massarwa, R., and Hanna, J.H. (2015). Establishing the human naive pluripotent state. *Curr Opin Genet Dev* 34, 35-45.
- Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol Cell* 32, 503-518.
- Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature* 469, 343-349.
- Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-560.
- Morin-Kensicki, E.M., Faust, C., LaMantia, C., and Magnuson, T. (2001). Cell and tissue requirements for the gene *eed* during mouse gastrulation and organogenesis. *Genesis* 31, 142-146.
- Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* 111, 197-208.
- Muller, J., and Verrijzer, P. (2009). Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr Opin Genet Dev* 19, 150-158.
- Nekrasov, M., Wild, B., and Muller, J. (2005). Nucleosome binding and histone methyltransferase activity of *Drosophila* PRC2. *EMBO Rep* 6, 348-353.
- Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* 4, 487-492.
- Nostro, M.C., Sarangi, F., Ogawa, S., Holtzinger, A., Corneo, B., Li, X., Micallef, S.J., Park, I.H., Basford, C., Wheeler, M.B., *et al.* (2011). Stage-specific signaling through TGFbeta family members and WNT regulates patterning and pancreatic specification of human pluripotent stem cells. *Development* 138, 861-871.

- O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene *Ezh2* is required for early mouse development. *Mol Cell Biol* *21*, 4330-4336.
- Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* *1*, 299-312.
- Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein *Suz12* is required for embryonic stem cell differentiation. *Mol Cell Biol* *27*, 3769-3779.
- Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). *Suz12* is essential for mouse development and for *EZH2* histone methyltransferase activity. *EMBO J* *23*, 4061-4071.
- Pasini, D., and Di Croce, L. (2016). Emerging roles for Polycomb proteins in cancer. *Curr Opin Genet Dev* *36*, 50-58.
- Pasini, D., Malatesta, M., Jung, H.R., Walfridsson, J., Willer, A., Olsson, L., Skotte, J., Wutz, A., Porse, B., Jensen, O.N., *et al.* (2010). Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res* *38*, 4958-4969.
- Pauklin, S., and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. *Cell* *155*:135-147.
- Pauklin, S., and Vallier, L. (2015). Activin/Nodal signalling in stem cells. *Development* *142*, 607-619.
- Pauklin, S., Madrigal, P., Bertero, A., and Vallier, L. (2016). Initiation of stem cell differentiation involves cell cycle-dependent regulation of developmental genes by Cyclin D. *Genes Dev* *30*:421-433.
- Pietersen, A.M., and van Lohuizen, M. (2008). Stem cell regulation by polycomb repressors: postponing commitment. *Curr Opin Cell Biol* *20*, 201-207.

- Riising, E.M., Comet, I., Leblanc, B., Wu, X., Johansen, J.V., and Helin, K. (2014). Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol Cell* 55, 347-360.
- Rossant, J. (2015). Mouse and human blastocyst-derived stem cells: vive les differences. *Development* 142, 9-12.
- Sauvageau, M., and Sauvageau, G. (2010). Polycomb group proteins: multi-faceted regulators of somatic stem cells and cancer. *Cell Stem Cell* 7, 299-313.
- Schoenfelder, S., Sugar, R., Dimond, A., Javierre, B.M., Armstrong, H., Mifsud, B., Dimitrova, E., Matheson, L., Tavares-Cadete, F., Furlan-Magaril, M., *et al.* (2015). Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat Genet* 47, 1179-1186.
- Schuettengruber, B., and Cavalli, G. (2009). Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. *Development* 136, 3531-3542.
- Seguin, C.A., Draper, J.S., Nagy, A., and Rossant, J. (2008). Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* 3, 182-195.
- Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* 32, 491-502.
- Surface, L.E., Thornton, S.R., and Boyer, L.A. (2010). Polycomb group proteins set the stage for early lineage commitment. *Cell Stem Cell* 7, 288-298.
- Tie, F., Stratton, C.A., Kurzhals, R.L., and Harte, P.J. (2007). The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. *Mol Cell Biol* 27, 2014-2026.

- Varambally, S., Dhanasekaran, S.M., Zhou, M., Barrette, T.R., Kumar-Sinha, C., Sanda, M.G., Ghosh, D., Pienta, K.J., Sewalt, R.G., Otte, A.P., *et al.* (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* *419*, 624-629.
- Wamaitha, S.E., del Valle, I., Cho, L.T., Wei, Y., Fogarty, N.M., Blakeley, P., Sherwood, R.I., Ji, H., and Niakan, K.K. (2015). Gata6 potently initiates reprogramming of pluripotent and differentiated cells to extraembryonic endoderm stem cells. *Genes Dev* *29*, 1239-1255.
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* *431*, 873-878.
- Wang, W., Lin, C., Lu, D., Ning, Z., Cox, T., Melvin, D., Wang, X., Bradley, A., and Liu, P. (2008). Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* *105*, 9290-9295.
- Xie, W., Schultz, M.D., Lister, R., Hou, Z., Rajagopal, N., Ray, P., Whitaker, J.W., Tian, S., Hawkins, R.D., Leung, D., *et al.* (2013). Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell* *153*, 1134-1148.
- Xu, C.R., Li, L.C., Donahue, G., Ying, L., Zhang, Y.W., Gadue, P., and Zaret, K.S. (2014). Dynamics of genomic H3K27me3 domains and role of EZH2 during pancreatic endocrine specification. *EMBO J* *33*, 2157-2170.
- Yusa, K., Zhou, L., Li, M.A., Bradley, A., and Craig, N.L. (2011). A hyperactive piggyBac transposase for mammalian applications. *Proc Natl Acad Sci U S A* *108*, 1531-1536.
- Zhao, X.D., Han, X., Chew, J.L., Liu, J., Chiu, K.P., Choo, A., Orlov, Y.L., Sung, W.K., Shahab, A., Kuznetsov, V.A., *et al.* (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* *1*, 286-298.
- Zylicz, J.J., Dietmann, S., Gunesdogan, U., Hackett, J.A., Cougot, D., Lee, C., and Surani, M.A. (2015). Chromatin dynamics and the role of G9a in gene regulation and enhancer silencing during early mouse development. *Elife* *4*.

## Figure legends

### Figure 1. Targeted deletion of *EZH2* in hESC

(A) Overview of *EZH2* structure and targeting strategy. Exons encoding CXC and SET domains are indicated. The gRNA sequence is underlined and protospacer adjacent motif highlighted in red. DNA sequence of the deletions in one *EZH2*<sup>+/-</sup> ESC line and one *EZH2*<sup>-/-</sup> ESC line is shown for both alleles. Mutation causes frameshift and premature stop codon. An additional line is shown in Figures S1 and S2.

(B) mRNA expression levels from RNA-Seq data revealing *EZH2*, *EED* and *SUZ12* transcript levels in *EZH2*<sup>+/+</sup>, *EZH2*<sup>+/-</sup>, *EZH2*<sup>-/-</sup> and *EZH2*<sup>-/-</sup> +*EZH2* ESC. Data show mean ± s.d.; n=3 biological replicates.

(C) Immunofluorescent microscopy of colonies from *EZH2*<sup>-/-</sup> ESC and control ESC. This analysis reveals a strong reduction in EZH2 levels in *EZH2*<sup>-/-</sup> ESC. The antibody was raised against a C-terminal epitope of EZH2; similar results were obtained using an alternative antibody raised against the N-terminal of EZH2 (Figure S2E). OCT4 expression indicates undifferentiated cells within a hESC colony. Arrows point to MEF. Scale bars, 100µm.

(D) EZH2, SUZ12 and the main isoform of EED are undetectable in *EZH2*<sup>-/-</sup> ESC by Western blot analysis, and are restored upon expression of a DOX-induced *EZH2* transgene. β-ACTIN is the loading control. Mass in kDa.

(E) H3K27me3 and H3K27me2 levels are reduced to background levels, and H3K27me1 levels are partially reduced, in *EZH2*<sup>-/-</sup> ESC. OCT4 expression in inset indicates undifferentiated ESC within the field of view. Arrows point to MEF. Scale bars, 100µm.

### Figure 2. *EZH2*-deficiency in hESC results in loss of H3K27me3

(A) Quantitative trend plot of H3K27me3 normalised ChIP-Seq reads over gene body  $\pm$  5kb. High CpG (HCP), intermediate CpG (ICP) and low CpG (LCP) promoters are shown separately.

(B) Scatter plot of H3K27me3 (x-axis) and H3K4me3 (y-axis) normalised ChIP-Seq reads in *EZH2*<sup>-/-</sup> relative to *EZH2*<sup>+/+</sup> (left) and relative to *EZH2*<sup>-/-</sup> + *EZH2* (centre), and *EZH2*<sup>-/-</sup> + *EZH2* versus *EZH2*<sup>+/+</sup> (right). All TSS shown in grey; TSS that are positive for H3K27me3 in *EZH2*<sup>+/+</sup> ESC highlighted in blue. Disruption of *EZH2* leads to a strong reduction in H3K27me3 levels at TSS, with little effect on H3K4me3 levels. Expression of a DOX-mediated *EZH2* transgene in the *EZH2*-deficient cells causes restoration of H3K27me3 levels to levels equivalent to *EZH2*<sup>+/+</sup>.

(C) ChIP-Seq tracks of *HOXB* (left) and *HOXD* (right) loci illustrate the loss of H3K27me3 in *EZH2*<sup>-/-</sup> ESC compared to control ESC. H3K4me3 is relatively unaffected. All ChIP-Seq data represent the average of three biological replicates for each cell line. These results were confirmed independently by qPCR analysis of ChIP DNA at several gene promoters (Figure S3C).

**Figure 3. Genes encoding developmental regulators are transcriptionally derepressed in *EZH2*-deficient hESC**

(A) RNA-Seq heatmap for *EZH2*<sup>-/-</sup> ESC and control ESC (three biological replicates per line). Shown are all differentially expressed genes between *EZH2*<sup>-/-</sup> and *EZH2*<sup>-/-</sup> + *EZH2* ESC.

(B) Top GO terms of differentially expressed gene sets. Numbers of genes are shown; example genes within each GO category are listed (right). Corrected *p*-values were calculated using a modified Fisher's exact test followed by Bonferroni's multiple comparison test.

(C) Gene set enrichment analysis of PRC2 targets (n=1299; defined by high *EZH2* and H3K27me3 promoter-localised ChIP-Seq values in *EZH2*<sup>+/+</sup> ESC) in genes that have been ranked according to their fold-change in transcription between *EZH2*<sup>-/-</sup> ESC and *EZH2*<sup>-/-</sup> + *EZH2*



ESC. The positive enrichment score (ES) reveals that genes selectively derepressed in the absence of *EZH2* are enriched in PRC2 targets ( $p < 0.001$ ; Kolmogorov-Smirnov statistic).

(D) Genes within the upregulated category have higher levels of promoter-localised *EZH2* (upper) and H3K27me3 (lower) in *EZH2*<sup>+/+</sup> ESC compared to an expression-matched set of genes and to downregulated genes. Data were compared using a Kruskal-Wallis test followed by Dunn's multiple comparison test.

(E) Genes encoding developmental regulators are transcriptionally derepressed in *EZH2*-deficient ESC. A subset of direct *EZH2* target genes is depicted as family groups. The colour of each circle represents the log<sub>2</sub> fold change in *EZH2*<sup>-/-</sup> ESC relative to *EZH2*<sup>-/-</sup>+*EZH2* ESC. The size of each circle represents the expression value of the gene in *EZH2*<sup>-/-</sup> cells. A similar pattern of target gene derepression is observed when comparing *EZH2*<sup>-/-</sup> ESC with *EZH2*<sup>+/+</sup> ESC (Figure S4C).

(F) ChIP-Seq and mRNA-Seq tracks of four genes encoding key developmental regulators illustrate the association between loss of H3K27me3 and transcriptional upregulation in *EZH2*<sup>-/-</sup> ESC.

**Figure 4. Transcriptional derepression occurs predominantly within discrete lineage-specific programmes**

(A) Single-cell RNA-Seq expression levels for six example PRC2-target genes in *EZH2*<sup>-/-</sup> ESC and *EZH2*<sup>+/+</sup> ESC, where each dot represent the results from a single cell. A pluripotency gene (*LIN28*) and housekeeping gene (*HMBS*) are shown for comparison. Robust upregulation of PRC2-target genes occurs in a subset of *EZH2*<sup>-/-</sup> ESC.

(B) Heatmap of single-cell RNA-Seq expression for *EZH2*<sup>-/-</sup> ESC (right) and *EZH2*<sup>+/+</sup> ESC (left).

Each column represents an individual cell. Each row represents an individual gene, grouped into three clusters corresponding to endoderm, mesoderm and ectoderm cell lineages. Shown are PRC2-target genes from within the hESC Scorecard assay, which is an assay that can classify differentiated cell lineages (Bock et al., 2011). Subsets of cells (boxed in purple) tend to mis-express many genes from within one lineage, but rarely mis-express multiple genes derived from more than one lineage.

**Figure 5. *EZH2*-deficient hESC are compromised in self-renewal and proliferation**

(A) Phase contrast images show representative colonies of *EZH2*<sup>+/+</sup>, *EZH2*<sup>-/-</sup> and *EZH2*<sup>-/-</sup> +*EZH2* ESC lines. Note the variable morphology of *EZH2*-deficient colonies. Scale bars, 100µm.

(B) *EZH2*<sup>-/-</sup> ESC show reduced ESC colony formation when plated as single SSEA4-positive cells at low density (6000 cells seeded per well). Data show mean ± s.d.; n=3 biological replicates. Data were compared using a one-way ANOVA followed by Bonferroni's multiple comparison test (\*\*\*, p<0.0005). Representative AP staining is shown underneath.

(C) *EZH2*<sup>-/-</sup> ESC have reduced capacity to self-renew when plated at clonal density. ESC colonies were categorised as undifferentiated, mixed or differentiated based on AP activity; examples shown underneath. Data show mean ± s.d.; n=3 biological replicates. Over 150 colonies were scored for each cell line. Data were compared using a one-way ANOVA followed by Bonferroni's multiple comparison test (\*\*\*, p<0.0005; \*\*, p<0.005; \*, p<0.05). Scale bars, 100µm.

(D) Immunofluorescent microscopy for OCT4 (undifferentiated marker) and SOX17 (early differentiation marker) reveals an increased prevalence for mixed and fully differentiated

colonies in *EZH2*<sup>-/-</sup> compared to control ESC lines. Representative images are shown underneath. Over 100 colonies were scored for each cell line. Scale bars, 100µm.

(E) Growth curve over 16 days reveals a significant proliferation defect in *EZH2*<sup>-/-</sup> ESC compared to control ESC. Data show mean ± s.d.; n=3 biological replicates. Data were compared between *EZH2*<sup>-/-</sup> ESC and each control ESC line using one-way ANOVA followed by Bonferroni's multiple comparison test (\*; p<0.05 for each comparison).

(F) Mitotic index was calculated for each ESC line by dividing the number of H3S10ph-positive cells by the total number of cells within a colony. The analysis was restricted to undifferentiated colonies (determined by OCT4 expression) of similar size in order to control for potential differences in cell state. Over 1000 cells were scored for each cell line. Data show mean ± s.d.; n=3 biological replicates. Data were compared using a one-way ANOVA followed by Bonferroni's multiple comparison test (\*\*\*, p<0.0005). Scale bars, 100µm.

**Figure 6. *EZH2*-deficient hESC can initiate differentiation, but are severely impaired in generating mature cell types**

(A) *EZH2*<sup>-/-</sup> ESC fail to generate teratomas. Mass of teratoma and kidney samples for indicated ESC lines, with images shown underneath (scale bar, 5mm). Additional images and histology analysis are provided in Figures S6A,B.

(B) ESC were induced to differentiate with retinoic acid for five days. Cell counts (upper) and crystal violet stain (lower) reveal that few *EZH2*<sup>-/-</sup> ESC remain after five days compared to control ESC. Short bars indicate mean values for the two biological replicates.

(C) *EZH2*-deficient ESC can generate early endoderm cells. Upper panel shows cell counts over endoderm differentiation timecourse. Lower panel shows flow cytometry analysis of endoderm markers KIT/CXCR4 in undifferentiated *EZH2*<sup>+/+</sup> ESC (black), day 5 endoderm

differentiated *EZH2*<sup>+/+</sup> (blue), undifferentiated *EZH2*<sup>-/-</sup> ESC (red), and day 5 endoderm differentiated *EZH2*<sup>-/-</sup> (purple). Inset numbers show percentage positive cells for each cell population (mean of three biological replicates, with range).

(D) *EZH2*-deficient ESC can generate early mesoderm cells. Upper panel shows cell counts over six days of mesoderm differentiation. Lower panel shows flow cytometry analysis of mesoderm markers KDR/PDGFR $\alpha$ .

(E) RT-qPCR analysis of endoderm, mesoderm, ectoderm and pluripotency genes in undifferentiated (black) and day 5 endoderm differentiated (blue) *EZH2*<sup>+/+</sup> ESC, and undifferentiated (red) and day 5 endoderm differentiated (purple) *EZH2*<sup>-/-</sup> ESC. Note that *POU5F1* and *NANOG* are also associated with ESC differentiation (Loh and Lim, 2011), which may underlie their elevated expression patterns in *EZH2*<sup>-/-</sup> ESC.

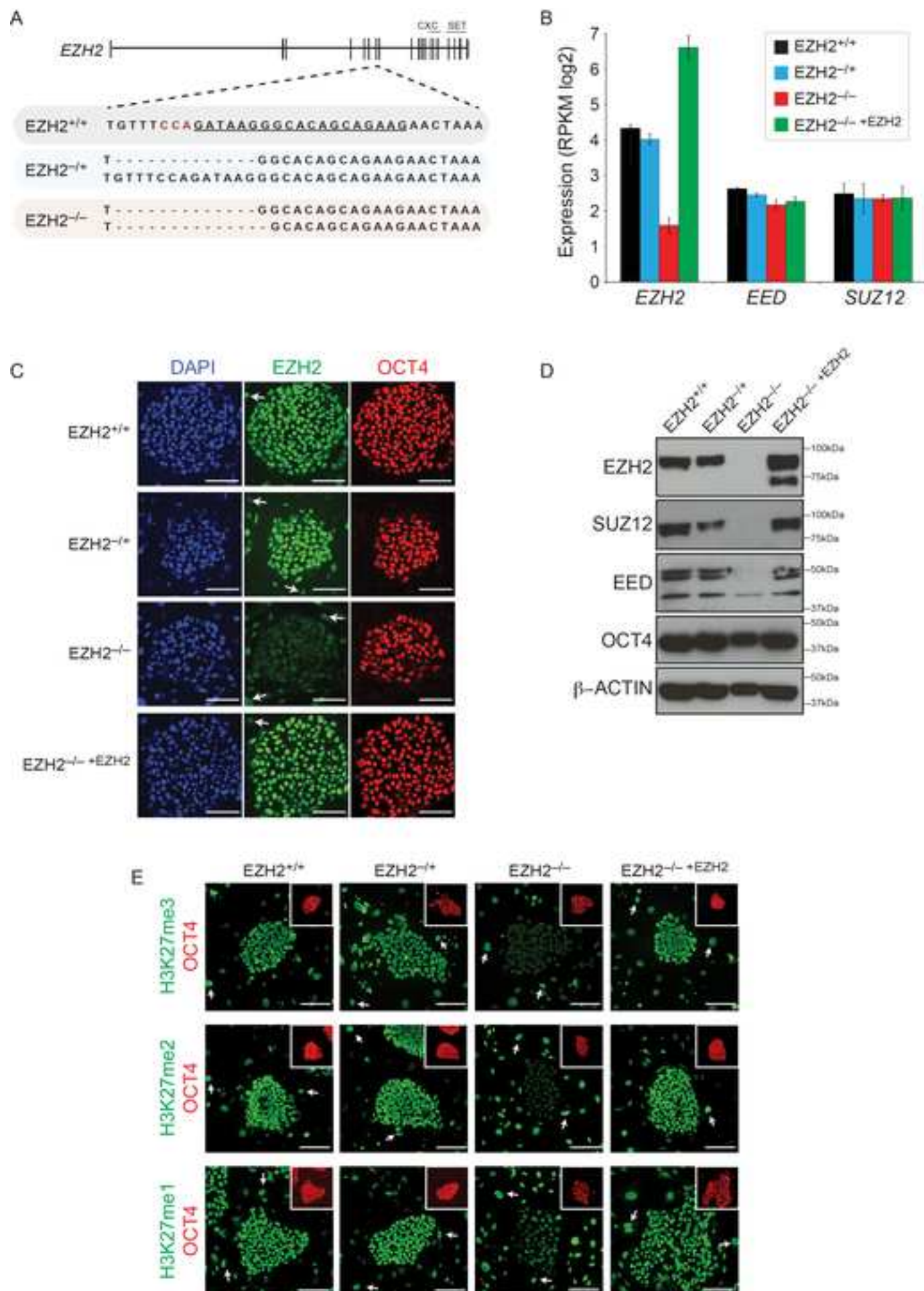
(F) qRT-PCR analysis of undifferentiated and day 2 mesoderm differentiated ESC.

(G) *EZH2*-deficient ESC can generate early ectoderm cells, but with significantly reduced efficiency compared to *EZH2*<sup>+/+</sup> ESC. Upper panel shows cell counts over 10 days of ectoderm differentiation. Lower panel shows flow cytometry analysis of ectoderm marker CD56 and undifferentiated ESC marker CD326. Note the significantly decreased efficiency of ectoderm differentiation in *EZH2*<sup>-/-</sup> ESC compared to *EZH2*<sup>+/+</sup> ESC ( $p=0.01$ ; unpaired two-sided  $t$ -test).

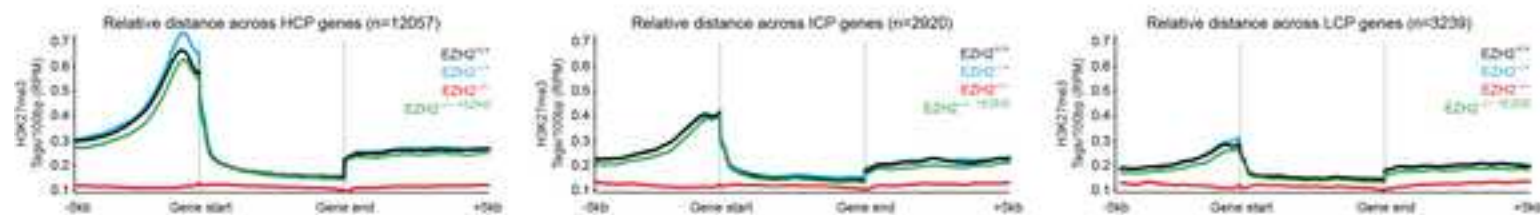
(H) RT-qPCR analysis of undifferentiated and day 10 ectoderm differentiated ESC. For all panels, data show mean  $\pm$  s.e.m of 3 biological replicates and were compared using an unpaired two-sided  $t$ -test (\*,  $p<0.05$ ).

**Figure 7. Proposed model summarising the role of *EZH2* in regulating transcriptional programmes and cell differentiation in hESC**

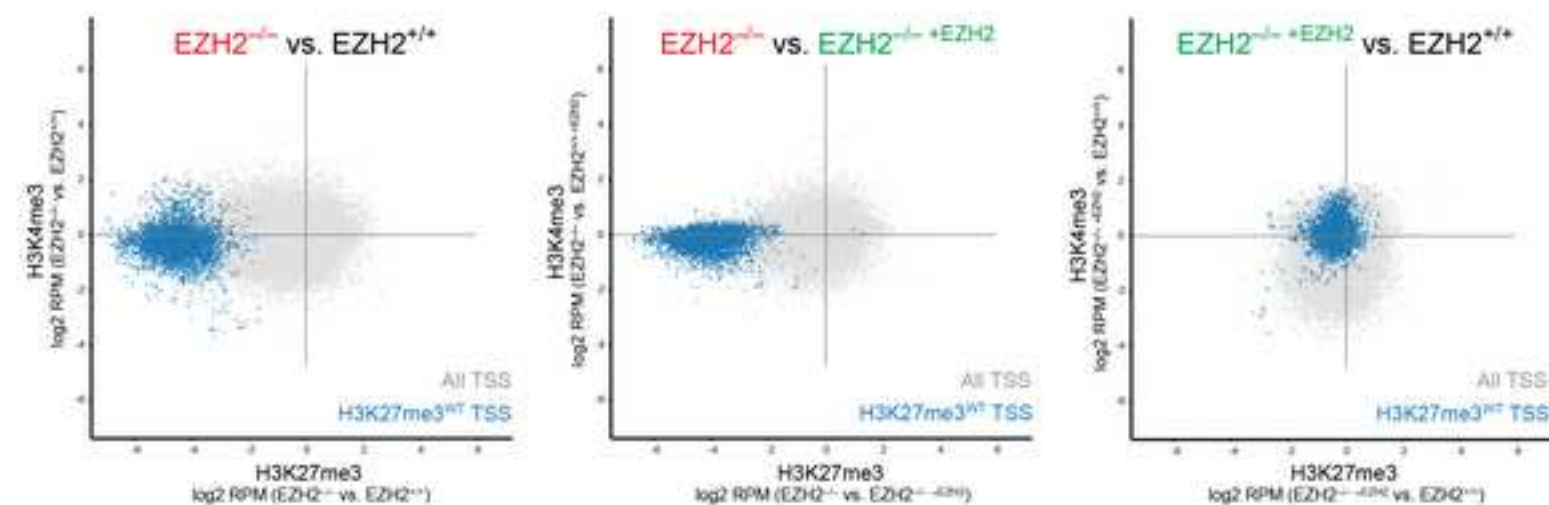
*EZH2*<sup>+/+</sup>, above; and *EZH2*<sup>-/-</sup>, below. *EZH2*-deficiency leads to loss of PRC2, transcriptional derepression of developmental regulators and self-renewal defects in hESC. Substantial cell loss (red crosses) and gene mis-regulation is observed upon differentiation of *EZH2*<sup>-/-</sup> ESC (e.g. to endoderm in example shown).



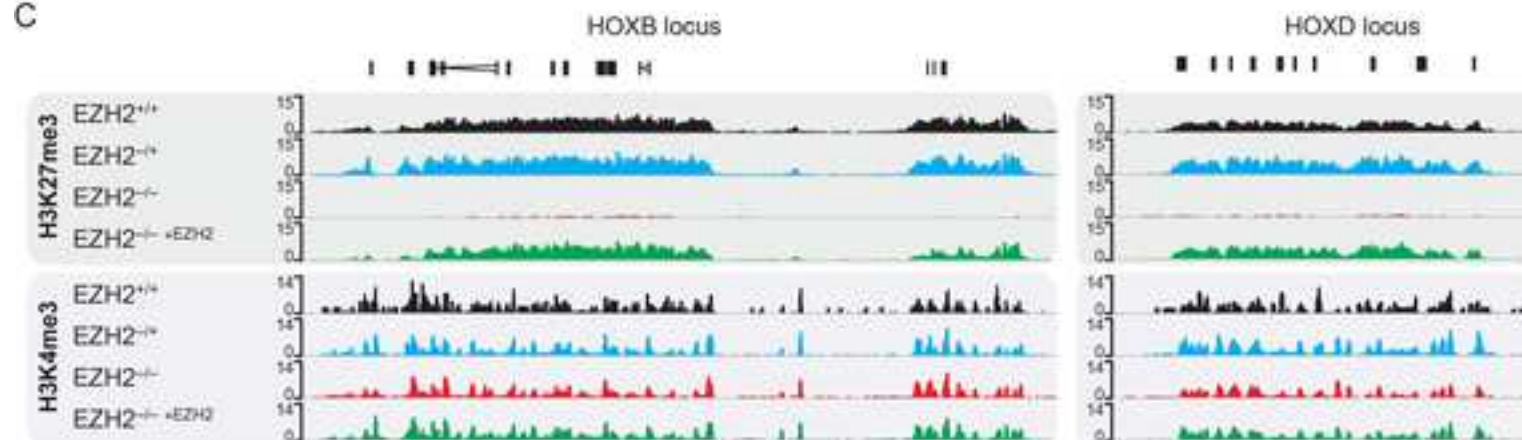
A



B

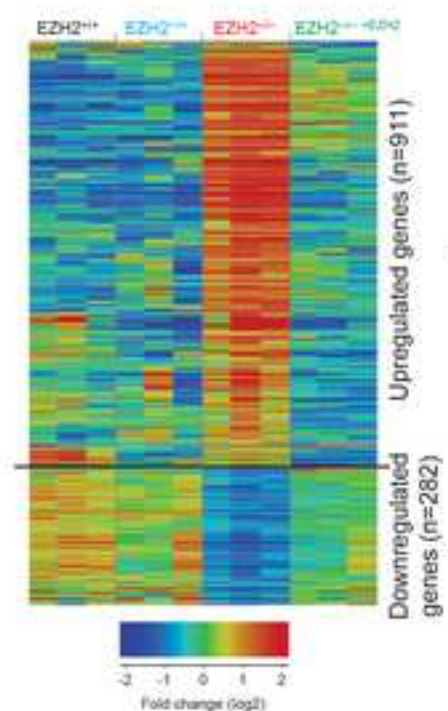


C



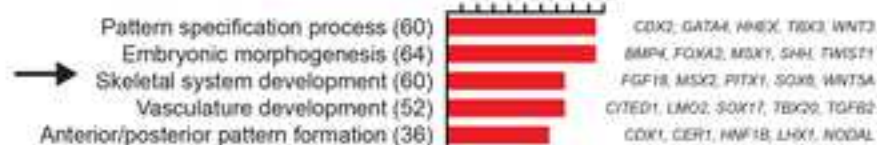


## A Differentially expressed genes



## B

## Gene Ontology

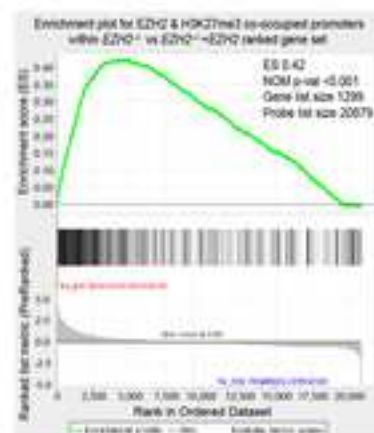


## Gene Ontology

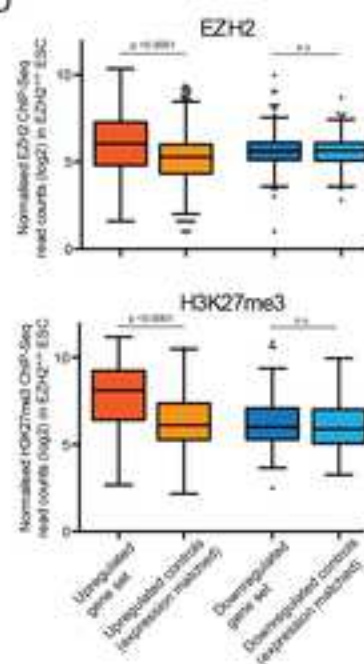
M phase (14)

Key genes associated with M phase include: RUB1, CDC7, DMC1, KIF15, MSH5.

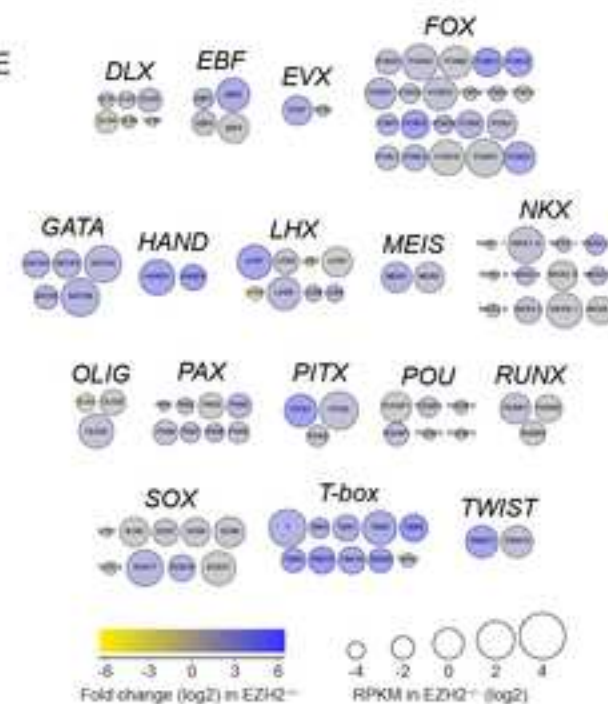
## C



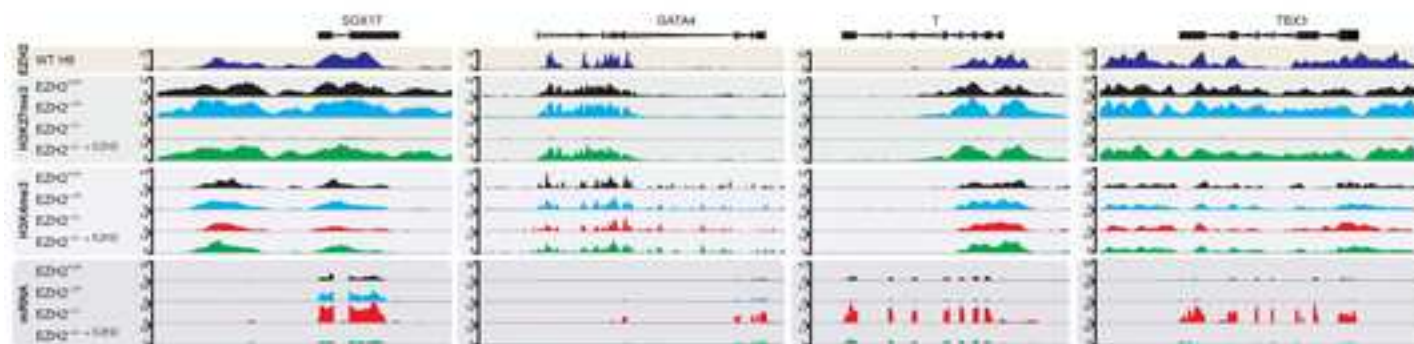
## D



## E



## F





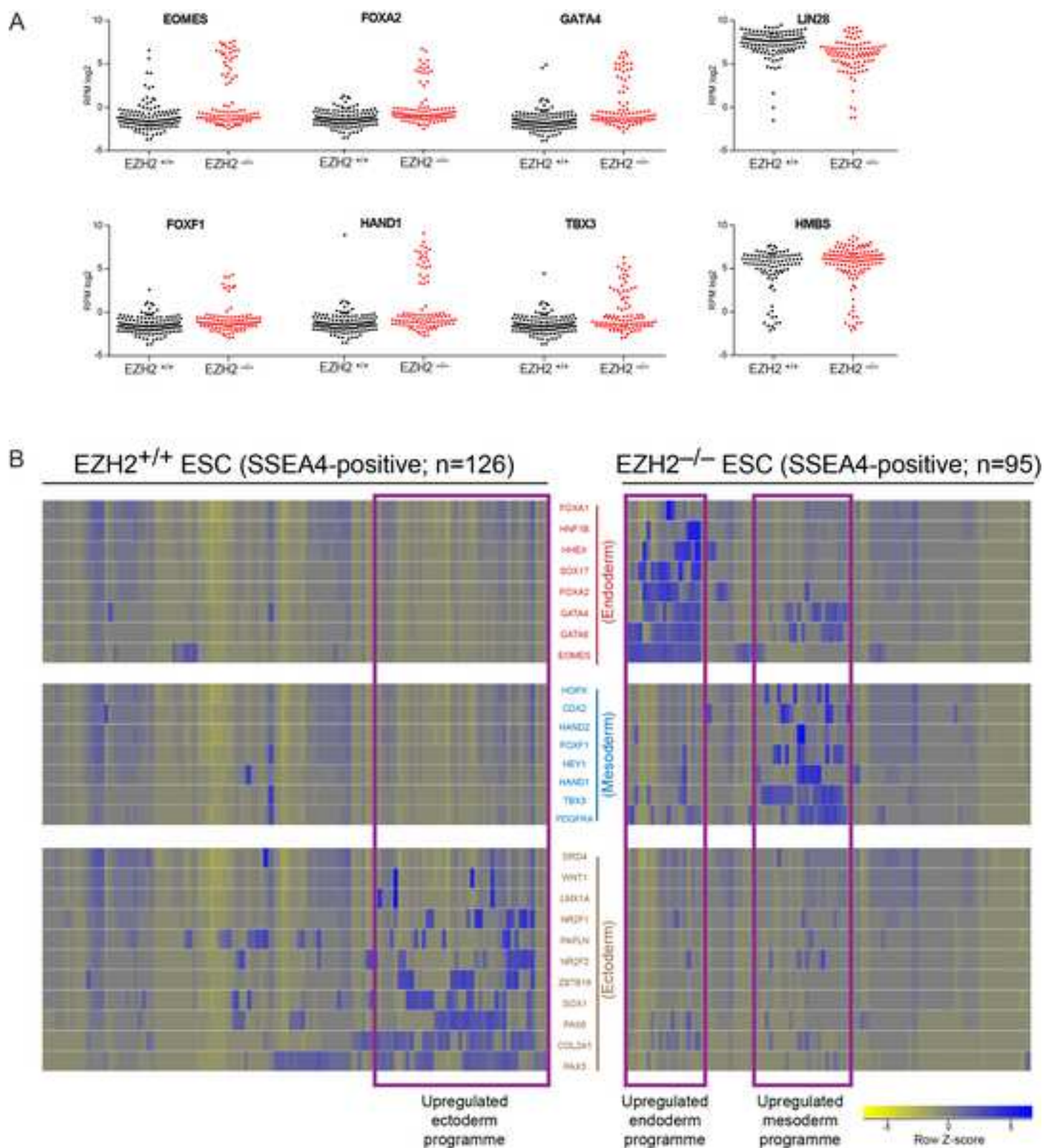
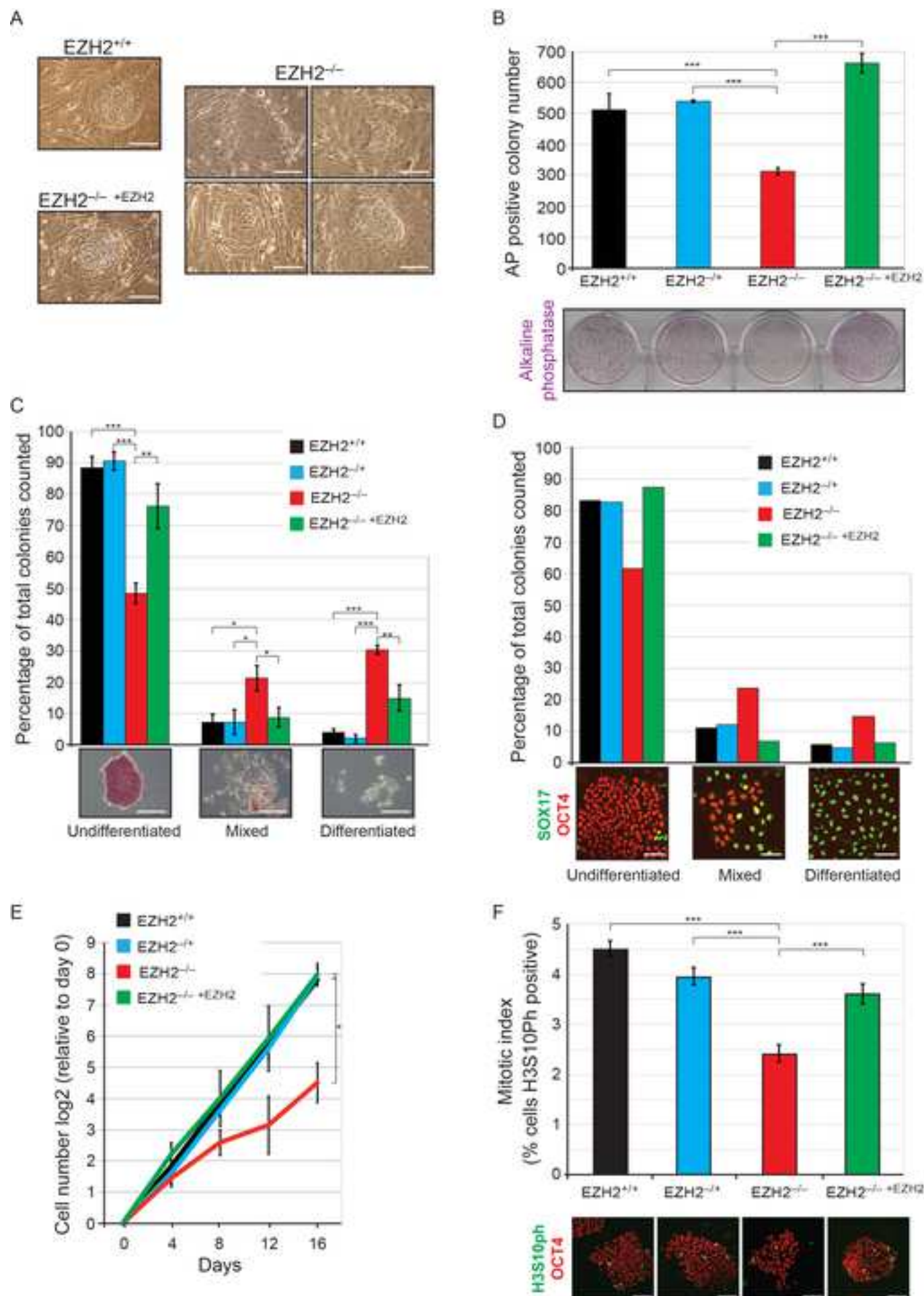


Figure 5

[Click here to download Figure Collinson\\_Fig5.tif](#)



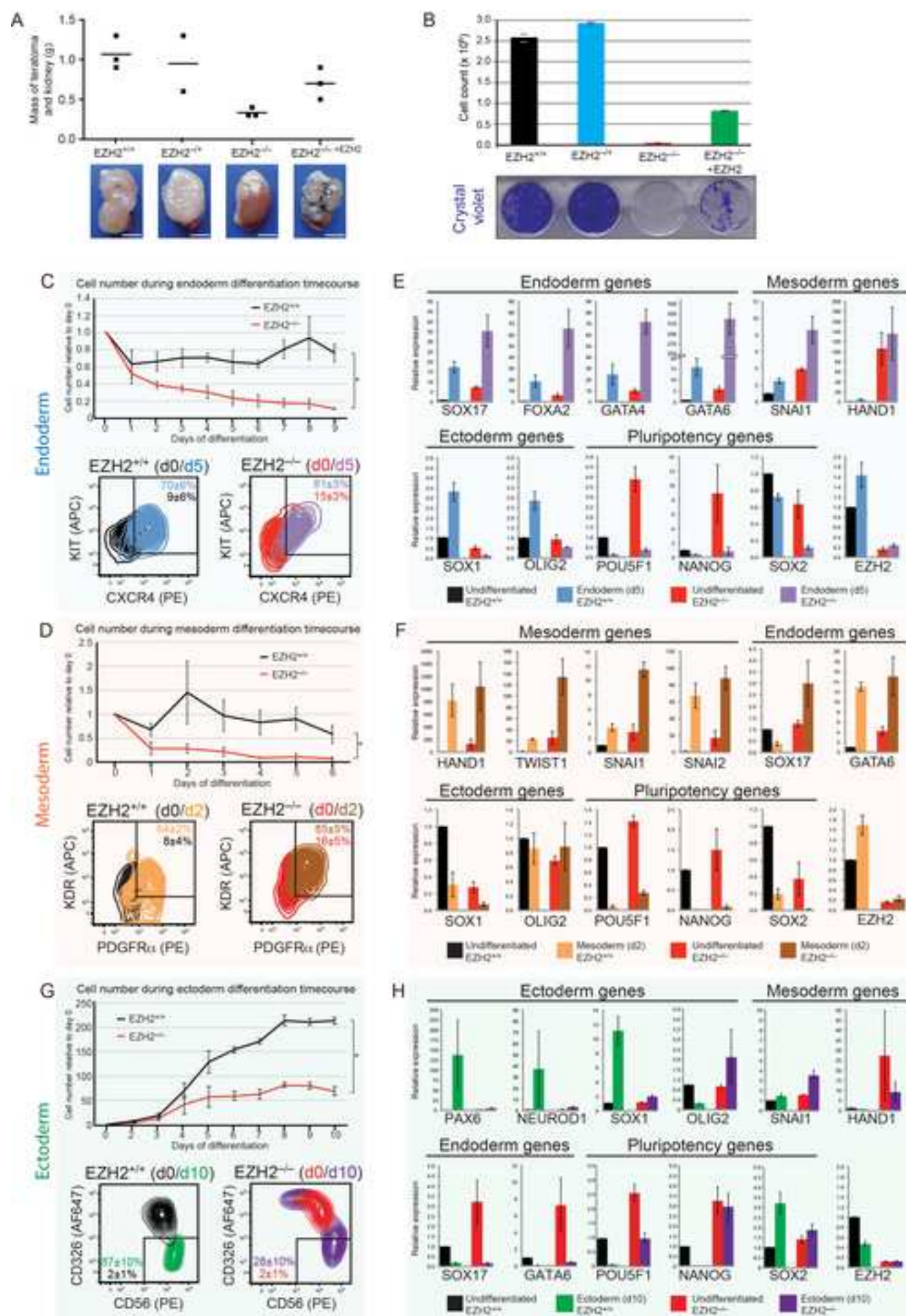


Figure 7

